ARI Research Note 97-18

Optimizing the Long-Term Retention of Skills: Structural and Analytic Approaches to Skill Maintenance

Alice F. Healy
University of Colorado

Research and Advanced Concepts Office
Michael Drillings, Chief

August 1997

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

University of Colorado

Technical review by

Michael Drillings

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED (from. . . to)
1997, August	Annual	May 1992-April 1993

4. TITLE AND SUBTITLE	5a. CONTRACT OR GRANT NUMBER
Optimizing the Long-Term Retention of Skills: Structural and Analytic Approaches to Skill Maintenance	MDA903-90-K-0066
	5b. PROGRAM ELEMENT NUMBER 0601102A
6. AUTHOR(S)	5c. PROJECT NUMBER B74F
Alice F. Healy (University of Colorado)	5d. TASK NUMBER 2901
	5e. WORK UNIT NUMBER C07

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
University of Colorado Department of Psychology Campus Box 345 Boulder, CO 80309-0345	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. MONITOR ACRONYM
U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-BR 5001 Eisenhower Avenue Alexandria, VA 22333-5600	ARI
	11. MONITOR REPORT NUMBER Research Note 97-18

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

COR: George Lawton

14. ABSTRACT *(Maximum 200 words)*:

Progress has been made on the topics of tank gunner skills, Morse code reception, color naming, instrument panel scanning, tests of the procedural reinstatement framework, mental calculation, memory for instances of categories, target detection, data entry, aspects of memory for lists, aspects of memory for course schedules, and vocabulary retention.

15. SUBJECT TERMS

Long-term retention; Analytic approach; Procedural reinstatement; Skill maintenance; Structural approach

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited	26	

Executive Summary of Progress: May 1, 1992 - April 30, 1993

OPTIMIZING THE LONG-TERM RETENTION OF SKILLS: STRUCTURAL AND ANALYTIC APPROACHES TO SKILL MAINTENANCE

MDA 903-90-K-0066 5/90 - 10/93
University of Colorado
Alice F. Healy, K. Anders Ericsson, and Lyle E. Bourne, Jr.

SCIENTIFIC OBJECTIVES: This research program seeks to identify the characteristics of knowledge and skill which are most resistant to decay due to disuse. The general goal is to elucidate principles which will specify those aspects of a complex skill that resist decay over periods of disuse and how they are distinguishable from more fragile components.

APPROACH: Four features of our program make it unique: First, our assumption has been that it is more crucial to optimize performance after a delay interval than to optimize performance during acquisition. Second, relative to most other empirical programs, we use longer retention intervals. Third, we have chosen to conduct experiments investigating a wide range of different skills and paradigms. Fourth, we have often used nontraditional methods to study retention.

PROGRESS: This research has led to the support or identification of several general principles about improving long-term retention or durability of skills. We focus on three classes of principles and illustrations of them from experimental investigations:

The first class of guidelines concerns ways to optimize retention through conditions of training. We discuss three general guidelines in this class. First, superior memory results from the use of cognitive procedures during learning. The procedural reinstatement framework is used to account for the observed superiority of memory for spatial order found in our studies of the retention of course schedule information. Second, retention is aided by prior familiarity. Memory for spatial information of course schedules was improved when the information could be related to previous experience. Third, learning is facilitated by distinctiveness of the information, as was evident with the spatial information in our study of list learning.

The second class of guidelines concerns ways to optimize the learning strategies used. We found in our study of mental arithmetic that the strategy used by the subject importantly influences retention and that a direct retrieval strategy leads to faster responding than does a strategy based on the use of an algorithm. Our study of vocabulary acquisition demonstrates that a direct retrieval strategy may also be achieved in that domain, but mediating associations may still exert an influence even when retrieval appears to be direct.

The last class of guidelines concerns ways to optimize memory through conditions of retention testing. In our study of vocabulary acquisition we saw remarkable recovery of retrieval speed after an initial retrieval. Hence, it appears that the use of a refresher or practice test before the critical test may have a profound impact on retention performance.

Some of our work also demonstrates the specificity of improvement in performance. Training on specific colors showed limited transfer to new colors in the Stroop color-word interference task. Although our original goal in this research program had been limited to an examination of the optimization of long-term retention, we have learned that optimizing retention does not guarantee generalizability. Our goals have broadened, so we are now exploring conditions of training, strategy utilization, and retention that simultaneously maximize both generalizability and long-term durability.

POTENTIAL APPLICATIONS: The overriding practical question of this research is how to ensure, through training, that a skilled worker (such as a code recipient, a tank gunner, or an aircraft pilot) has a behavioral tool kit which is just as or nearly as permanently functionable as his or her hardware kit. The eventual goal is to be able to make relevant recommendations about training routines for long-term skill maintenance.

SELECTED NEW GENERAL ARCHIVAL PUBLICATIONS:

Fendrich, D.W., Healy, A.F., & Bourne, L.E. (in press). Mental arithmetic: Training and retention of multiplication skill. In C. Izawa (Ed.), *Cognitive psychology applied*. Hillsdale, New Jersey: Erlbaum.

Healy, A.F., Clawson, D.M., McNamara, D.S., Marmie, W.R., Schneider, V.I., Rickard, T.C., Crutcher, R.J., King, C., Ericsson, K.A., & Bourne, L.E., Jr. (in press). The long-term retention of knowledge and skills. In D. Medin (Ed.), *The psychology of learning and motivation*. New York: Academic Press.

Healy, A.F., Fendrich, D.W., Crutcher, R.J., Wittman, W.T., Gesi, A.T., Ericsson, K.A., & Bourne, L.E., Jr. (1992). The long-term retention of skills. In A.F. Healy, S.M. Kosslyn, & R.M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes, Volume 2* (pp. 87-118). Hillsdale, NJ: Erlbaum.

Annual Report for the period May 1, 1992 to April 30, 1993
Optimizing the Long-term Retention of Skills:
Structural and Analytic Approaches to Skill Maintenance
Principal Investigator: Alice F. Healy
Co-Investigators: K. Anders Ericsson and Lyle E. Bourne, Jr.
Postdoctoral Research Associates: Robert Crutcher, Cheri King, Danielle McNamara, Vivian Schneider, and Grant Sinclair
Graduate Research Assistants: Deborah Clawson, Oliver Hammerle, Rajan Mahadevan, William Marmie, and Tim Rickard
Undergraduate Research Assistants: Mary Jensen, Rachel Moore, Gregory Rully, Ursula Williams

General comments. We are in the process of refocusing our research directions to meet the Army's needs in the area of foreign language training. Although we continue to work on long-term retention and although some of this work is relevant to issues concerning foreign language training, we have formulated a series of new investigations which are aimed directly at improving the training of foreign languages. We submitted a formal proposal to the ARI on this topic at the end of August and learned in March that this project will be funded.

We were invited to write a new chapter summarizing our research on long-term skill retention for the series Psychology of Learning and Motivation. We completed the process of adapting the report we presented in Orlando, Florida, at the 1992 ARI contractors' meeting to make it suitable as a book chapter. We submitted this chapter to the editor of the series, Douglas Medin, who had a very favorable reaction to it. We then made minor revisions to the chapter in response to some optional suggested changes from Dr. Medin. This chapter is now in press (SEE APPENDIX A).

Anders Ericsson and Rajan Mahadevan left the University of Colorado at the end of December to begin new positions at Florida State University, Tallahassee, Florida.

We prepared and presented a report for the ARI contractors' meeting held at the beginning of February in Alexandria, Virginia.

Lyle Bourne has written a chapter entitled, "History of research on thinking" for the Handbook of Perception and Cognition, Volume 12 (Thinking and Problem Solving), edited by Robert Sternberg. He co-authored this chapter with Roger Dominowski. The chapter is now in press. Similarly, Alice Healy was invited to write a chapter entitled, "The long-term residue of training and instruction" for the Handbook of Perception and Cognition, Volume 10 (Memory), edited by Elizabeth and Robert Bjork. She will be co-authoring this chapter with Grant Sinclair.

Meetings and visits. Ursula Williams, an advanced undergraduate student from the University of South Carolina, visited our laboratory in the summer of 1992. She was a volunteer research assistant who helped us with a number of different ongoing studies.

On May 12, Lyle Bourne delivered an invited colloquium at the

University of Tubingen, West Germany. The talk was entitled, "Becoming an expert: Experiments and theory of cognitive automatization."

In June, Liang Tao presented a paper with Alice Healy as co-author at the First International Conference on Chinese Linguistics held in Singapore. The paper was entitled, "Cognitive strategies in discourse processing: A comparison of Chinese and English speakers."

In June, Thomas Cunningham of St. Lawrence University visited us for ten days to work on collaborative research concerning recall of order information.

David Fendrich visited us in August, and we worked with him on a manuscript summarizing our collaborative research on data entry including a new study recently completed in Dr. Fendrich's laboratory at Widener University.

David Rubin visited us in September and delivered an invited colloquium in the Psychology Department on autobiographical memory. Prior to his visit, we studied an article in which Dr. Rubin presents a new analytic technique that examines the rank order of items on a memory test as a way to determine whether, despite forgetting, subjects exhibit the same pattern of responses at different delay intervals. In this article Rubin argues that there is no change in the pattern of recall responses as a function of delay, suggesting that it may not be necessary to study long-term retention. We contrasted this argument with the recent article by Schmidt and Bjork pointing to studies of contextual interference which demonstrate that the methods used to optimize performance during learning do not necessarily optimize performance during retention so that it is crucial to study long-term retention. Our meeting with Rubin was, thus, stimulating and productive.

In November, four of us (Alice Healy, Lyle Bourne, Anders Ericsson, and Tim Rickard) attended the annual meeting of the Psychonomic Society in St. Louis, Missouri. Anders Ericsson co-authored a poster with Robert Crutcher, now of the University of Illinois at Chicago, on mediation processes in memory retrieval. Lyle Bourne presented a paper co-authored with Tim Rickard on mental multiplication. Dr. Healy attended meetings of the Governing Board of the Psychonomic Society, the Publications Committee of the Psychonomic Society, the New Journal Planning Subcommittee of the Publications Committee of the Psychonomic Society, and the Editorial Board of the journal Memory & Cognition. Drs. Healy and Bourne also attended the meeting of the Executive Committee of Division 3 (Experimental Psychology) of the American Psychological Association.

Alice Healy, Vivian Schneider, and Deborah Clawson visited Colorado State University in November to attend the doctoral dissertation defense by Cheri King. Alice Healy chaired the defense as her primary advisor.

Antoinette Gesi visited us in December, and we worked with her

on a manuscript summarizing our collaborative research on data entry.

At the end of December, Anders Ericsson made a trip to London, UK, where he presented a paper at a CIBA symposium and at a meeting of the Royal Society of Medicine. He also made a presentation at a press conference organized by the CIBA Foundation. The title of his presentations was "Can we create gifted people?"

In February, Alice Healy and Lyle Bourne presented a report summarizing our research at the ARI contractors' meeting in Alexandria, Virginia.

In March, Anders Ericsson presented a paper at the Seventh Annual Florida Conference on Sensation, Perception, Cognition, and Action in Tampa, Florida, on the topic of "Individual Differences in Expert Performance: The Role of Motivation, Talent, and Practice."

In April, Deborah Clawson, Vivian Schneider, and Alice Healy attended the joint meeting of the Rocky Mountain Psychological Association (RMPA) and Western Psychological Association in Phoenix, Arizona. Deborah Clawson presented a paper describing her work on Morse Code reception. The title of the paper was "Acquisition and Retention of Morse Code Reception: Part-Whole Training." The co-authors were Deborah Clawson, Alice Healy, Anders Ericsson, and Lyle Bourne. Alice Healy attended the Executive Committee Meeting of the RMPA, as the new President Elect of the RMPA.

In April, Anders Ericsson presented an invited colloquium at the Department of Psychology, Institute of Cognitive Science, McGill University, in Montreal, Canada, on the topic of "Expert Performance: Its Structure and Acquisition."

At the end of April, Dr. Janet Proctor of Purdue University presented a paper at the Midwestern Psychological Association in Chicago, Illinois. The paper was entitled, "The effects of practice on the word frequency disadvantage in letter detection." The co-authors were Janet Proctor and Alice Healy.

Tank gunner skills. Reaction-time analyses for the long-term retention experiment in our series on tank gunner skills have been completed. The analyses used the average of each session half for each subject's first 14 sessions. One reaction-time analysis looked at the time to make an identification of a target, and another analysis looked at the time to fire after making an identification. Because performance reached asymptote so quickly, no main effects of session, session half, or their interaction emerged from either analysis. Whereas there are noticable trends towards a decrease in fire and identification times as a function of session, and indeed, an increase after the first 6-month retention interval, none of these differences were large enough to make them statistically reliable. This failure to attain statistical significance is most likely due to the relatively small number of subjects (i.e., 6) tested in this experiment. Essentially the same testing materials and procedure were used in the part/whole experiment (which included 24 subjects) with statistically significant results emerging from the analyses in

that study.

Data from 5 of the original 6 subjects were collected in a 15th session conducted approximately 22 months after the end of the initial training phase. However, because of simulator malfunctions, data from only 3 of the 5 subjects could be used.

We have completed the analysis of the reaction time data through the 22nd-month retention test (15th session) of the long-term training experiment. As before, there does not appear to be any substantial loss of skill, in terms of reaction time, across the delay, with even a trend towards slight improvement in the time to make an identification.

Additional planned analyses were conducted which compared response times from each consecutive pair of sessions in the first experiment. These analyses confirmed that for both the time to make an identification and the time to fire measures, no reliable difference across the two-week retention interval was present.

In the autoslew experiment, we found that during training the time taken to fire on a target was significantly reduced for subjects who trained on one part of the task at a time. However, several subjects reported dislike of this condition. For example, when asked whether they thought that being trained by being moved by the commander helped them or not, one subject responded: "I felt comfortable moving the tank by myself. If anything, it frustrated me and made me sloppy." Other subjects felt comfortable however, and said things like: "Yes, it let [me] concentrate on lining up the tank, and not having to find it."

In summary, although no overall deficit was observed on the measure that reflected the searching component (time to make an identification) as a result of the part-training, significant improvement during training was observed on the measure that reflected the time required to fire on a target. This component is, arguably, more important.

Finally, we continued our work on a manuscript describing our series of experiments on the long-term retention of the tank gunner skill. Additionally, we have continued our review of the literature on the long-term retention of complex skills. We plan to submit this manuscript to a professional journal in applied psychology.

Bill Marmie is the primary investigator on this project, with Alice Healy as the faculty advisor.

Morse code reception. We are preparing a manuscript reporting our studies of the acquisition and retention of Morse code reception skills. We plan to submit that manuscript to a research journal in applied psychology.

The opposing results of Morse code reception part-whole training and tank gunner part-whole training (as we reported in the 1992 ARI contractors' meeting in Orlando) led to a further set of analyses of

our Morse code study. In our Morse code part-whole training study, we had found that those subjects who initially learned the difficult subset of code-letter pairs did not gain an advantage over those subjects who learned the entire set right from the start of training. In contrast, in our tank gunner part-whole training study, we did find an advantage for those subjects who first learned the more difficult of two serial subtasks over those who first learned the whole task. We attributed this difference in results to a difference in degree of mastery of the difficult task during initial training. The difficult tank gunner subtask could be mastered during the training sessions, whereas the difficult Morse code subtask could not be fully mastered during the available training time. It is possible, however, that the difference in advantage was due to a difference in type of part task used. The part task in the Morse code study was learning part of the stimulus set, but in the tank gunner study the part task was learning one step in a serial process.

We sought to explore the effect of such "part-process" training on Morse code acquisition by a new analysis of data from a previous Morse code experiment. In that experiment, groups of subjects were trained on one of three tasks: the whole Morse code reception task (which we called "code-to-letter"), the part task of segmenting the auditorily presented code into its elements ("code-to-dida"), or the part task of translating the pattern of elements into the appropriate letter ("dida-to-letter"). Subjects were trained for two sessions, then they returned two weeks later for a retention test on the task they had learned. It was after this retention test that the relevant data were obtained. At this time, all subjects were presented with one session of training on the whole code-to-letter task; for those in the two part-task groups, this was their first training on the whole task. After training, all subjects were tested on the whole task. Accuracy on the whole task tended to be highest for those subjects who had received whole-task training throughout the experiment. The two groups initially trained on part tasks did not differ in accuracy, performing at levels just over 60%. We also examined reaction times on the whole task for the different training groups. Although the accuracy measures did not yield significant differences among the groups, reaction time measures did show reliable differences. Specifically, when performing the whole task the code-to-dida group, which trained initially on segmenting the auditory signal into its dots and dashes, was significantly slower than the other two groups. Thus, "part-process" training did not yield an advantage over whole-task training. Two caveats should be noted, however. First, the relevant training and testing were conducted after the retention interval rather than following the initial training as in the tank gunner experiment. Second, it is unclear whether one of the part tasks could be labelled as "more difficult" than the other; although the task of segmenting the heard code showed less improvement with practice than did the second subtask, the two groups had virtually identical accuracy levels on their initial pretests.

To explore the possibility that initial training on one of the part tasks influenced later learning of the whole task, we have completed analyses of the errors made by the three training groups on

the whole-task test described above. Overall, the error profiles of the three groups were similar. There were only two differences in the profiles. First, the code-to-dida group was the only group that did not commit significantly more same-length errors than would be expected by chance on two-element codes. Interestingly, the other part-task group, dida-to-letter, had shown a similar lack of significance during initial training, suggesting that in this aspect the code-to-dida group's performance changed in response to learning the other part of the whole task. The second difference was that on three-element codes the dida-to-letter group was the only one that committed significantly fewer two-different errors than chance. The source of this difference is less clear, because both of the part-task groups had exhibited this tendency during initial training.

A presentation describing the part-whole analyses and error analyses of Morse code was made at the joint Rocky Mountain Psychological Association/Western Psychological Association conference.

Deborah Clawson is the primary investigator on this project. Her Masters thesis summarized this research; she received her Masters degree in May. Her primary faculty advisor is Alice Healy, with Anders Ericsson and Lyle Bourne as her secondary faculty advisors.

Color naming. Extensive training and testing of subjects in our primary experiment has been completed -- 12 hours of training under either the interference condition or the noninterference condition, as well as pretests, posttests, and one-month retention tests.

In this experiment there were three groups of subjects. Control subjects received no training; experimental subjects received either the interference condition or the noninterference condition in 12 hours of training. All subjects completed pretests, posttests, and one-month retention tests. Substantial speed-up on a number of subtests was apparent at the posttest, with some forgetting during the retention period. Analyses of the tests involving orthographic manipulations did not reveal any lessening of Stroop interference when the visual configurations of the words were altered. The tests of reading speed for black and white words did not suggest that reading speed decreased after extensive Stroop training. By far the most interesting finding was that the effects of training on the Stroop-related tests were different depending on the color set used in the tests. Subjects were trained extensively on one of two color sets (half of the subjects in each condition were trained on one set, and the remaining subjects were trained on the other set) but were tested on both color sets. In an analysis of reaction times on the pretest and posttest, it was found that responding became significantly faster after training. The specificity of training was apparent in that reaction times decreased more on tests using the trained color set than on tests using the untrained color set. An analysis of the posttest and retention test further described the specificity. In this analysis there were no significant effects or interactions involving test time, and there was no significant forgetting over the retention interval. However, subjects were faster on the trained set than on the untrained set on the two types

of tests -- color patch naming and Stroop color naming -- that required responding with the color name, but not on the two types of tests that required reading the color word. This advantage for the trained set on tests requiring color naming was clear only for the group trained on the Stroop color naming task. The specificity of training to the trained color set was especially striking given the lack of significant improvement during the actual training sessions.

Item analyses for the Stroop-trained subjects in our primary experiment yielded further suggestive results about the specificity of automaticity training. Specifically, reaction time performance was examined for each combination of color and word that was used in training (e.g., the word red printed in green letters, the word purple written in green letters, or the word green written in red letters). Both subjects' reaction times revealed that the interaction between word and color was highly significant by the end of training. This interaction suggests that the effects of training are to some degree specific at the level of word-color combinations, rather than solely at the level of words or colors.

Protocol analyses of the first subject's practice performance yielded a puzzling result. In the first session of training, the subject gave two main types of protocols, one in which she reported registering the color immediately and another in which she reported first registering the word then registering the color. As would be expected, her reaction times were faster on trials for which she reported the first type of protocol. As training progressed, however, her reports of this faster access disappeared, with virtually all trials yielding the second type of protocol. We hope that analysis of the other subject's protocols will allow us to explore whether such performance reflects an explicit suboptimal strategy choice or is for some subjects an unavoidable result of exposure to the stimuli.

Further lines of investigation into the specificity of Stroop training effects are being planned for Deborah Clawson's dissertation research, including using training and test materials with different shades of the same colors, to see whether specificity is related to the response term or to the stimulus color. Additional studies would include training subjects on one set of colors and color words then testing them on a set of incongruous stimuli using the same colors but different words, or vice versa.

The Cognitive Science Society has accepted a paper on our specificity findings for their conference in Boulder (SEE APPENDIX B). This paper will be published in the Annual Proceedings of the Cognitive Science Society.

This research is currently being conducted by Deborah Clawson and Cheri King, with Alice Healy and Anders Ericsson as primary advisors.

Instrument panel scanning. We completed a draft of a manuscript combining the two experiments from Vivian Schneider's dissertation with the related experiment conducted more recently. We plan to

submit this manuscript for publication in a professional research journal.

We are currently designing new experiments on the topic of instrument panel scanning. For example, in one planned experiment the subjects are cued as to which configuration is relevant on a given trial rather than which rule; that is, the cuing is done by row rather than by column.

We have administered a long-term retention test to all the subjects we were able to contact who were previously tested in the spring semester of last year. It was hoped that this test would allow us to determine if the random practice group retained its advantage over the blocked group after a nine- to ten-month delay. Unfortunately, we were unable to get enough subjects in one of the groups to make a strong comparison although we found the trend to be in the same direction as before (i.e., in favor of the random practice group). We also conducted an analysis comparing the original test last spring and the retention test in the fall in order to examine forgetting. We found no forgetting evident for the measure of accuracy; in fact there was no difference between the retention test at one or four weeks and the retention test at nine or ten months. For the measure of reaction time, we obtained a significant difference between the speed of reponses on the retention test after one or four weeks and the retention test after nine or ten months. However, the advantage was for the test after the longer retention interval. That is, subjects were significantly faster after nine or ten months than they had been on the earlier retention test. This was found to be true even for the first block of the two tests, and in fact the subjects were faster on the first block after nine or ten months than they had been on the last block of the previous retention test. Thus, it seems that subjects continue to improve in this task even over a nine or ten month retention interval.

Vivian Schneider is the primary investigator on this research, with Alice Healy as the primary faculty advisor and Anders Ericsson and Lyle Bourne as the secondary faculty advisors.

Tests of the procedural reinstatement framework. The chapter summarizing our earlier work on this topic was published this year in one of the two volumes edited by Alice Healy, Stephen Kosslyn, and Richard Shiffrin in honor of William K. Estes.

Danielle McNamara successfully completed and defended her dissertation entitled "The advantages of generating extended to skill acquisition and retention: Procedural implications." (SEE APPENDIX C). This thesis included three experiments in which the generation effect was extended to skill learning and the acquisition and long-term retention of facts stored in semantic memory. In these experiments subjects were trained in either a read or generate condition. The first experiment examined the generation effect with both simple and difficult multiplication problems. As expected, a generation effect does occur for difficult multiplication problems, and not with the simple multiplication problems. The second

experiment examined the long-term memorial consequences of generating versus reading nonwords. Those subjects who had generated the nonwords during training had significantly better long-term retention of the nonwords than did those who had simply copied the pairs. In addition, it was found that within the read condition, those subjects who had spontaneously utilized mnemonic strategies to learn and retrieve the nonwords had significantly better recall than did subjects who had not. The findings of these two studies are understood in terms of theories of proceduralization. It is hypothesized that subjects in a generate condition are more likely to develop procedural skills during training than are subjects in a read condition. These skills, developed during training, are subsequently reinstated at test. We believe that the significant factor in the generation effect is that subjects develop cognitive procedures to learn and recall the test items. Both of the studies described above support this hypothesis. In the third experiment, an overall generation advantage was not found to extend to the elementary school setting for the initial learning of multiplication skills, although generating did lead to more efficient retrieval than did reading on a retention test after a one-week delay.

A report of the first two experiments described above was resubmitted to a professional journal for publication.

Two further experiments were conducted which examined the generation effect with addition and multiplication problems. In the first experiment, problem type was manipulated as a within-subjects variable (i.e., subjects saw both addition and multiplication problems). In the second experiment, problem type was manipulated as a between-subjects variable (i.e., subjects saw only addition or only multiplication problems). In both experiments, subjects read aloud the problems and either read or generated the answers. After a brief interpolated task, subjects' memory for the answers was tested, first with a recall procedure and then with a recognition procedure. According to our hypothesis that the important factor for the generation effect to occur is that connections be activated, a generation effect should only occur for the multiplication problems and not for the addition problems because for multiplication problems there exists a limited number of ways of obtaining a product, whereas in addition, there are numerous ways of obtaining a sum. Therefore, the "connections" are more important for the process of multiplication than for addition. As predicted, when problem type was manipulated between-subjects, a generation effect occurred only for the multiplication problems and not for the addition problems. However, we found that a generation effect occurred with both types of problems when problem type (addition vs. multiplication) was manipulated within-subjects.

A related experiment was previously completed (in collaboration with David McCain, as his undergraduate honors thesis) which investigated the generation effect with simple and difficult multiplication problems. Subjects read aloud problems and either read or generated the answers. After a brief interpolated task, subjects' memory for the answers was tested, first with a recall procedure and then with a recognition procedure. Our procedural

account led to the prediction of a larger generation effect for simple than for difficult multiplication problems in this task which tests episodic memory rather than skill acquisition because the subjects would be likely to have stronger memorial links between the operands and the products for the simple than for the difficult problems. However, our findings showed a generation effect for both types of problems.

Our findings of differential generation effects when problem type was varied between subjects versus within subjects when comparing addition and multiplication problems led to the prediction that differential generation effects may also be found when comparing simple and difficult multiplication problems. Indeed, this year we completed a follow-up study which manipulated the type of multiplication problems (i.e., simple or difficult) as a between-subjects variable. As expected, a generation effect was found with only the simple multiplication problems and not for the difficult multiplication problems.

These four experiments together supported our hypothesis that the generation effect will only occur for arithmetic problems for which the problem and answer are highly associated. This finding supports a procedural account of the generation effect whereby the critical factor for the generation effect to occur is that cognitive procedures be activated. We are currently preparing a manuscript describing this series of four experiments.

These projects are being performed by Danielle McNamara with Alice Healy as the faculty advisor.

Our related collaboration on the generation effect with J. David Mason, a graduate student in the Accounting Department of the School of Business, continued this year. We began a new experiment which utilized similar materials to previous experiments in this research line. College students solved accounting problems with simple and complex numbers. In this study, however, we included a new between-subjects variable: early or late indication of the nature of the recall task. That is, we either informed or did not inform subjects that they would be required to remember only the first two digits of the problems. The experiment was conducted with the assistance of Ursula Williams.

Previously, our major finding involved a difference in recall for simple and complex sums in addition problems, with simple numbers being recalled better. In this new experiment, with the early instructions we were able to eliminate almost completely the simple-complex difference on the recall test. In contrast, however, the instructional manipulation had no effect on the results of the recognition test. That is, a simple-complex difference persisted for recognition irrespective of the instructional condition.

In addition, in this experiment, which involved memory for the answers (i.e., sums) of addition problems, there were three processing conditions: read, verify-addend, and verify-sum. We are in the process of formulating different hypotheses about the failure

to obtain a generation effect (i.e., an advantage for the verify-sum condition) in this experiment. Results of the study conducted by Danielle McNamara comparing recall of addition versus multiplication problems have raised a question pertinent to this study. Perhaps the procedural connections between addends and sums are only formed when multiplication problems as well as addition problems are being performed by the subject in a single session. This would amount to a type of procedural-transfer process occurring during encoding. According to this reasoning, a generation effect simply may not exist for addition problems on their own. Alternatively, the lack of time to develop a mnemonic link in the verify-sum condition in the present study may have hampered performance relative to that in the read and verify-addend conditions. At least one subject in the present study reported having more time to rehearse items in the read condition.

Work on the manuscript describing the results of our series of experiments continues. We have focused attention on the important role memory for rounded and exact numbers plays in many everyday and applied memory situations.

We explored the question of whether the persistence of a simple-complex difference (henceforth to be called a rounded-exact difference) that we obtained for recognition memory in the third experiment was an artifact of the task factor. With the assistance of Rachel Moore, an undergraduate student, we initiated a fourth experiment, which removed the task factor from the experiment. Preliminary results suggest that the rounded-exact difference does not in fact persist in the recognition test with the simpler task.

The primary investigators on this research are J. David Mason and Bill Marmie, with Alice Healy as the faculty advisor.

Mental calculation. We completed the write-up of a manuscript summarizing Tim Rickard's Masters thesis research on mental arithmetic. This manuscript was submitted for publication to a major professional journal. We received positive editorial feedback with suggestions for revision, and we are now in the process of completing the suggested changes.

Two new mental arithmetic studies were completed this year. These studies are extensions of our previous explorations of practice and transfer of arithmetic skill. In the first study, subjects were trained on basic multiplication and division problems, and then tested on the exact problems on which they practiced, as well as on various versions of the practice problems. One group was tested immediately after practice and after a one-month retention interval, and another group was tested only after the one-month delay. The overall goals of this study were (a) to understand in some detail the nature of the representation of arithmetic facts in memory, and (b) to explore further retention of any improvements in performance gained with practice. The results support a model of fact representation which we proposed in our earlier work: For each triplet of numbers defined by multiplication/division (e.g., 6,9,54), there are three unique abstract representations in memory, for example, (6,9)=54, (6,54)=9, and (9,54)=6, where the order and

perceptual characteristics of the elements in parentheses are only an artifact of the notation, and are not a part of the actual representation. Practice transfers positively and substantially across changes within one of these units of representation (e.g., across a change in operand order for multipication), but does not transfer between representational units (e.g., practice on 54/6 does not transfer to a subsequent test on 6 x 9). With respect to retention, our previous research showed excellent retention of skill for arithmetic problems on which subjects practiced. The new study largely confirmed this previous finding, although the degree of skill retention was not as great as in the previous studies. Tim Rickard and Lyle Bourne reported this study at the meeting of the Psychonomic Society in November.

The second study explored whether the problem size effect in mental multiplication (the fact that small problems such as 3 x 4 are solved more accurately and quickly than large problems such as 6 x 9) disappears after extensive practice. The character and extent of change in the problem size effect with practice should yield insight into the cognitive mechanisms that underlie this robust effect. Preliminary data analyses show that the problem size effect does diminish substantially with practice, and, for problems with products smaller than about 50, the effect essentially disappears. When problems with larger products (e.g., 6 x 9, 7 x 8) are considered, however, a problem size effect remains. Future research will be directed toward understanding the mechanisms underlying this residual effect, and exploring the extent to which the diminished problem size effect is retained over an extended retention interval.

We also have plans to pursue two additional arithmetic projects which we have previously described (a set of empirical studies exploring strategies used to solve arithmetic problems and a computer simulation model of arithmetic performance).

We began a new project which explores skill acquisition, transfer, and retention, using a pseudo-arithmetic task. In our first experiment, subjects were trained on a set of math problems which require a novel calculation algorithm, symbolized by the pound sign (#). Specifically, subjects were given two elements from a simple arithmetic progression, and were required to generate the third (next) element. The generic progression used was one in which the third element is the second element plus the difference between the first and second elements, plus 1. For example, the answer to 7 # 15 = _ is 15 + (15 - 7) + 1 = 24. A reverse algorithm was defined which yields the answer to a problem with the second element of the series missing (e.g., 7 # _ = 24). The # and = symbols were used to make the task visually similar to standard arithmetic tasks, such as multiplication and division. Subjects were trained for 5 sessions on 6 different problems with the third element missing (type I problems), and on 6 different problems with the second element missing (type II problems); they were then given both an immediate and a delayed test on the training problems (no-change problems), on training problems with the missing element changed (type change problems, e.g., type I problems became type II problems, and vice versa), and on new problems. On one-third of the trials subjects

were probed to determine whether they used the algorithm or retrieved the answer directly from memory.

The basic idea of this experiment was to study skill acquisition on a task similar to standard arithmetic which presumably would require execution of an algorithm early during practice, with a transition to direct retrieval of facts (i.e., the answers to specific problems) from memory as skill improves. In preliminary analyses, we have uncovered four results of primary interest. First, the strategy probing data showed that there was a complete transition to retrieval after about 60 blocks of practice. Second, the reaction time and standard deviation data did not conform to power functions, as predicted by all current models of skill acquisition. Rather, the power function appears to hold only when analyses are limited to exclusively algorithm or retrieval trials, but breaks down when all the data are analyzed together. We have formulated and are currently testing an alternative skill acquisition model which appears to be able to predict the observed deviations from the power function. Third, our previous work on arithmetic showed learning to be very specific. We found that practice on _ = 7 x 8, for example, not only does not transfer to new (unpracticed) problems at test (e.g., _ = 6 x 9), it also does not transfer to division problems that contain the same digits (e.g., 56 = _ x 8). In this new study, we found the same results; learning was very specific to the problems on which subjects practiced, and did not transfer to either type change or new problems. Finally, preliminary analysis of the retention data showed moderate retention of skill for no-change problems. There was a drop in the frequency with which the retrieval strategy was reported, from over 90% of the time on the posttest, to around 50% of the time on the retention test. Interestingly, however, when subjects reported using the retrieval strategy on the retention test, the reaction times were the same as those for trials on which subjects reported using retrieval on the posttest. Thus, the loss in skill appears to be completely describable in terms of a drop in the probability of using retrieval. Additional analyses are planned to explore this finding further.

Tim Rickard is the primary investigator of this research. His primary faculty advisor is Lyle Bourne, with Alice Healy as a secondary advisor.

Memory for instances of categories. We designed and conducted a set of studies to explore the acquisition and retention of conceptual knowledge for faces. In the original pilot study, subjects were shown a set of faces that had similar features (i.e., family faces), followed by a set of faces that were dissimilar from the faces in the first set and from each other (i.e., non-family faces). Subjects were then tested on selected faces from both of these sets, as well as on new family faces not previously seen. During the test, subjects were asked to identify faces that were old family faces. Preliminary results demonstrate that subjects are slightly more likely to identify family faces (both old and new) than non-family faces at test. There was no evidence for a tendency to identify old non-family faces more than new non-family faces. Results indicate that memory for the faces was strongly feature-driven, and that the

particular faces used in each phase of the experiment had a major impact on the results.

In follow-up research, we experimented with using non-overlapping feature pools for family and non-family faces in an effort to get more substantial differences across conditions. We also manipulated the salience level of the faces in each group (i.e., the degree to which a given face represents its group or family) in order to test the notion of feature-driven category representation. We also tested some subjects immediately after acquisition, and others after a one-day or two-week delay. We found evidence that, on the immediate test, subjects were able to discriminate family from non-family faces, but that after delays, subjects lost their memory for context (i.e., family or non-family) and based their responses more on familiarity than they did at immediate test. Thus, subjects mistakenly identified non-family faces as family faces more often at delay than at immediate test. This effect was more pronounced at the 2-week delay interval than at the 24-hour delay interval.

This result is interpreted as evidence for a change in the use of detail-oriented to familiarity-based memory processes which occur as a function of time since encoding. Longer response latencies for correct rejections and larger effects due to salience (i.e., featural overlap to a feature frequency category representation) in the delay conditions provided converging evidence in support of this conclusion.

Further research has begun which investigates the nature of the category representation more closely. In particular, we hope to ascertain the utility of an exemplar model in accounting for the results in this study. The methodology of this new study is virtually identical to that of the last study. However, now we manipulate subjects' exposure to family faces during the familiarization phase. One third of the subjects see all of the family exemplars an equal number of times. The familiarization phase for the remaining subjects includes repeated presentations of two of the exemplars randomly embedded in the set of family faces. For half of these subjects, the repeated exemplars are constructed from the most frequently occurring features in the family; the repeated exemplars for the remaining subjects are constructed from the least frequently occurring features in the family.

By using the sleeper effect (increased number of misclassifications of familiarized distractor faces at delays) as a diagnostic, we hope to determine if greater exemplar retrievability influences category inclusion judgments for new exemplars. If subjects are acquiring and/or accessing a category representation based on exemplars rather than an abstract representation, we expect fewer false alarms for non-family faces at delays for subjects in the repeated exemplar conditions than for subjects in the non-repeated exemplar condition.

Also by manipulating the representativeness of the repeated exemplars (close to the prototype vs. far from the prototype), we hope to see if the nature of the retrieved exemplars plays a role in

inclusion judgments. Our new design will enable us to relate our findings to the earlier work on concept attainment by William K. Estes and to the theoretical framework he proposed.

Oliver Hammerle and Tim Rickard are engaged in this research with Lyle Bourne as their primary faculty advisor.

Target detection. Vivian Schneider and Alice Healy submitted a manuscript to the professional journal Memory & Cognition summarizing six experiments on our investigation of phonetic factors influencing letter detection. This manuscript was accepted for publication and is now in press (SEE APPENDIX D).

In this manuscript, we concluded that visual unitization and word frequency, as well as phonetic factors (specifically a mismatch between the letter representation and the phonetic representation), contribute to the errors in letter detection. We found no evidence for the hypothesis that the location of the letter within the word is responsible for letter detection errors. We were able to show that two levels of visual processes are involved in letter detection and have opposite effects. One type involves the well-formed Gestalt qualities of the letter o that make that letter easier to detect than the letter t and, thus, fewer letter detection errors are made on the letter o than on the letter t in words that contain both letters. This type of visual process operates at the letter level. We also found a second visual effect that involves differences in computer fonts and operates at the word level. Subjects made fewer detection errors on common words with fonts that are script-like relative to more usual fonts. This finding was predicted because the script-like fonts tend to interfere with the usual visual configuration of common words and, thus, word unitization. In the case of the visual process that operates on the letter level, if the letter is made less regular, then more errors are made in letter detection. On the other hand, at the word level, if the usual word configuration is distorted, then word unitization is interfered with and fewer letters are missed.

In the revisions we recently made to the manuscript, we now explain how the two levels of visual processes that are involved in letter detection (but have opposite effects) fit with the unitization hypothesis. The unitization hypothesis states that words and letters are processed in parallel, but once a word is identified processing stops even if all the component letters have not been identified. More common words such as the and of are processed as units before there is time to identify all the letters. In the case of the visual process that operates at the letter level, if the letter is made less regular, then more errors are made in letter detection. This is explained by noting that the more well-formed configuration of the letter o makes it more likely to be identified before the word has been identified and processing stops. On the other hand, at the word level, if the word is made less regular, then word unitization is disturbed and more time is required to process the word, thereby allowing the subject longer to identify the letters so fewer letters are missed.

Vivian Schneider is the primary investigator on this research, with Alice Healy as the primary faculty advisor.

Progress was also made this year in the write-up of two earlier studies on letter detection. One of these studies is being done by Alice Healy in collaboration with Sheldon Tetewsky, who is now of McGill University, and includes adaptations of the interactive-activation model to account for the word frequency disadvantage we find with the letter detection task. The second study is being conducted by Alice Healy in collaboration with Janet Proctor and includes four experiments exploring the effects of practice on the pattern of errors in the letter detection task. This year we also completed the data collection and analysis of a fifth experiment for this study, which was performed at Purdue University by Dr. Proctor. In addition, a paper summarizing this research was accepted for presentation at the Midwestern Psychological Association. Dr. Proctor presented this paper at the conference on April 30, 1993.

Data entry. Dr. David Fendrich of Widener University completed testing subjects in a new experiment in our continuing series on data entry. Later, he completed the data analysis for this experiment. We added this experiment to our revised manuscript reporting the earlier results of the experiment we conducted with Antoinette Gesi, now a graduate student at the University of California, Santa Cruz. Dr. Fendrich visited us this year, and we worked together during his visit on the revisions of our manuscript. Ms. Gesi visited us later this year, and we worked together during her visit on some new revisions of this manuscript. We recently submitted the revised manuscript to a professional journal for publication. The primary advisor on this research project has been Alice Healy, with Lyle Bourne as the secondary advisor.

Temporal, spatial, and item components of memory for lists. A new experiment in our series examining the role of distinctive items in short-term recall of order information has been completed by our colleague Thomas Cunningham at St. Lawrence University. This experiment, which utilized the same design as our initial von Restorff experiment, required subjects to generate a missing letter from the to-be-remembered letter segments. In June, Dr. Cunningham visited our laboratory and brought us the data from this experiment to analyze. Our project has provided many interesting findings. Three analyses, parallel to analyses conducted on the von Restorff experiment, have been completed and suggest that when a letter is missing from a segment, overall recall of that segment is hampered. But this effect is mediated by both the presentation rate and the instructional condition under which a subject operates. When subjects are instructed that the missing item should be recalled as well as the other items in a segment and subjects are in the slow presentation rate condition, they tend to recall just as many items in the item-missing segments as they do in the item-present segments. However, in all other conditions there is an advantage for item-present segments over item-missing segments. That is, the difference between recall of item-missing and item-present segments seems to be eliminated given sufficient instructions and time. This

result poses an interesting challenge to the perturbation model because of the influence of a motivational source (i.e., instruction to remember an item preferentially) which appears to interact with short- and long-term memorial processes. Even more interesting is the apparent necessity for that motivational source to be sufficiently aroused by the stimulus encoding and retrieval conditions. Notably, in our study of the von Restorff effect in this same paradigm, we found no effects of motivating instructions, and the stimulus encoding conditions involved (showing an item in red) presumably did not require any additional encoding operations on the part of the subject beyond the request of the experimenter. On the other hand, the recall of a missing item in the present study did require additional operations on the part of the subject beyond the experimenter's request (either at encoding or at retrieval). We are planning on exploring ways to modify the perturbation model to adapt to this unique situation.

In addition to collecting segment recall data, Dr. Cunningham and his students collected output-order data to provide us with more information on the effects of generation on output strategies in a short-term memory task. The output-order analysis results suggest that overall subjects responded from left to right only 70% of the time. We found further that when a segment contained a missing letter, there was only a 52% probability that the response would be output in a left-to-right order, but when a segment did not contain a missing letter, the likelihood of a left-to-right output order response was 89%. For the faster presentation rate the likelihood of recalling a segment in order increased with the length of the retention interval, but for the slower presentation rate there was a smaller trend in the opposite direction. Similarly, for the faster rate the likelihood of recalling a segment in order was greater for the more temporally distant first segment than for the second segment, but for the slower rate there was a smaller trend in the opposite direction.

Bill Marmie is engaged in this research with Alice Healy and Thomas Cunningham as his primary advisors.

Also during his visit, Thomas Cunningham worked with Alice Healy on the revision of a manuscript summarizing some of our earlier research on this topic. New data analyses and new model fits were completed for this revision. Subsequently the manuscript was provisionally accepted for publication in the journal Memory & Cognition, with suggestions for additional revisions. These revisions were completed, and the manuscript is now in press (SEE APPENDIX E).

Our investigation that directly compares short- and long-term retention of features of common objects has revealed some interesting findings. We conducted, with the assistance of Greg Rully, a full experiment which examined 32 subjects' memory for features of two United States coins, a familiar (i.e., modern) penny, and a unfamiliar (i.e., from 1941 and not used today) Liberty dime. Three analyses were conducted using both a strict and a lenient method of scoring. The lenient method scored an item (or feature) as correct

if it was simply included in the drawing of the coin. The strict scoring method required that the item was placed in the correct location on the drawing. Nine items were scored for each coin (e.g., the word "liberty" and the head on the front of the coins).

In our penny-first condition we were able to replicate virtually every difference between items obtained by Nickerson and Adams (1979) although our subjects tended to score a little bit worse on each item. This replication points to the reliability of our findings, and it allowed us to examine our new conditions with some confidence about the method and procedures we used. The item differences we found suggest that some features are more difficult to remember than others. For example, whereas more than 95% of subjects could remember that there was a head on the front of both the penny and the dime, fewer than 20% remembered the word "liberty" on the penny (and only 80% remembered it on the dime, which had been referred to as the "Liberty" dime). As expected, we found that when we tested subjects one week after study, recall was worse than on the immediate test, especially by the strict method of scoring.

Most crucially, we found that, overall, recall for the Liberty dime was better than that for the familiar penny regardless of the scoring method used. In fact, recall for the unfamiliar Liberty dime after the one-week retention interval was better than that for the familiar penny initially. This finding points out the strong advantage of intentional study (of the unfamiliar Liberty dime) relative to incidental learning (of the familiar penny), because the opportunities to study the familiar penny outside the laboratory were presumably extremely frequent, whereas only ONE-MINUTE was allowed for studying the unfamiliar Liberty dime. In addition there was a reliable interaction of penny-first/dime-first group and coin, with subjects who recalled the penny first performing better on the Liberty dime than subjects who recalled the dime first. This result is somewhat remarkable given that subjects never study a penny at all in either condition. This finding points out the strong advantage of an initial recall attempt, which may invoke some of the memory processes responsible for the facilitation caused by intentional learning. A similar finding pointing to a memory advantage of a recall attempt was made earlier in our laboratory in the context of short-term retention in the distractor paradigm (see Cunningham, Healy,' & Williams, 1984).

On a theoretical level, we emphasize an important, but previously overlooked, variable which is a critical factor in memory for coins: intentional study. Previous studies demonstrating exceptional retention of details in visual images (e.g., Shepard, 1967) have been contrasted with the Nickerson and Adams (1979) study, upon which the present investigation was based. This contrast led to the development of a basic empirical issue: Under what conditions are common objects remembered? That is, it is unclear why exceptional memory existed in one case but not the other. Our study convincingly demonstrates that intentional study significantly enhances long-term retention. Notably, even after a one-week retention interval, memory for an unfamiliar coin (an out-of-mint "liberty" dime), studied for only one minute, was superior to memory

for a U.S. penny, a coin people likely encounter almost every day.

We prepared a manuscript summarizing our findings on this project for submission to a professional journal. An exhaustive literature review was conducted identifying all research that cited the influential Nickerson and Adams (1979) study, which inspired the present experiment.

A shorter manuscript describing this study was submitted and accepted for a poster presentation at the 15th annual meeting of the Cognitive Science Society. The manuscript will be published in the Society's proceedings and is included here (SEE APPENDIX F).

Bill Marmie is preparing to present these findings at the Psychology Department's annual mini-convention. Several future directions are possible. First, to resolve convincingly the issue of whether visual images are well-retained or not, it would be desirable to replicate our experiment using a recognition task rather than a recall task. Second, the question of whether daily exposure to U.S. pennies makes a penny more memorable than an unfamiliar dime could be explored using a design which explores the lower limits of study time for both coins necessary for successful encoding. A study time manipulation (60, 45, 30 and 15 sec period) could be introduced to allow us to make this judgment. These studies may form the basis for Bill Marmie's doctoral dissertation.

This investigation is being conducted by Bill Marmie with Alice Healy as the primary advisor.

We also worked on our study comparing long-term memory for temporal, spatial, and item information about word lists. We conducted new analyses and extended our manuscript summarizing this study, which we plan to submit to a professional journal on memory. This work is being conducted by Grant Sinclair, with Alice Healy as the primary faculty advisor and Lyle Bourne as the secondary advisor.

Temporal, spatial, and item components of memory for course schedules. We conducted a study on this topic in collaboration with Cheri King, a former doctoral student at Colorado State University who performed this research as her doctoral dissertation. In this study, students from both the University of Colorado and Colorado State University participated as subjects and learned course schedules from either their own institution or from the unfamiliar campus. Each of the subjects learned a course schedule in one session, followed by retention tests (each consisting of a recall questionnaire, a class list, and a map test) at one week and two-, three-, or four-week intervals. Ninety-six subjects were tested. In accord with our procedural reinstatement framework, we expected superior memory for spatial information about course locations only for subjects familiar with the relevant campus.

For the training data, there was a significant effect of type of information during the initial trials and again during the fifth and sixth trials. For the questionnaire test and retest data, there was not a significant interaction of type of information by familiarity.

However, the trend was in the predicted direction. More detailed analyses were conducted showing that recall of spatial information was only superior for those subjects in the familiar condition.

Additional analyses were carried out showing significant forgetting from test to retest on the recall questionnaire, the class listing test, and the map test. On the same three tests, there was a significant difference among the four types of information, as well as a significant interaction between type of information and training orientation. As expected, on the class listing test there was a significant interaction between type of information and familiarity.

The initial write-up of this experiment was completed as the doctoral dissertation of Cheri King (SEE APPENDIX G), who succesfully defended her dissertation this year and received her doctoral degree from Colorado State University.

This project is being conducted by Cheri King, with Alice Healy as the primary advisor and Lyle Bourne as a secondary advisor.

This year Major William Wittman of the U.S. Air Force Academy and Alice Healy revised their manuscript describing earlier research on this topic and submitted it to a professional journal for publication.

Vocabulary retention. We carried out further analyses of our keyword experiments that looked at the effects of extended practice on retrieval of Spanish-English vocabulary pairs. In addition, we began working on various projects to present and/or publish our findings to date. We began working on a manuscript that discusses the findings from our earlier keyword studies as well as our more recent work on the effects of extended practice. We performed associated re-analyses of the data from those studies. We also prepared and presented a poster at the Psychonomics Society meeting in St. Louis in November.

A new study was planned; pilot subjects' testing was completed, and additional subjects are currently being tested. This study is motivated by our efforts to relate our training investigations to language learning outside the laboratory. In particular, we are interested in comparing the speed of retrieval of English equivalents. Extended training studies showed a speed-up in retrieval consistent with the power law. However, even after over a hundred blocks of training, the retrieval speed of our experimental subjects was slow (800-900 ms) compared to other estimates of retrieval speed, for example, in naming familiar objects. To explore asymptotic performance for this type of retrieval, we recruited groups of subjects differing in their mastery of Spanish, ranging from undergraduate majors to bilingual speakers. These subjects are being tested with the same Spanish-English items used in our earlier training studies. The retrieval performance of the Spanish-speaking groups will then be compared to that of the trained subjects taking into account the far greater experience of Spanish for the Spanish-speaking groups.

Our earlier reported studies of retention of Spanish vocabulary learning were based on a relatively small number of repetitions during learning. Reasonable levels of retention were observed after a one-week delay, but clear evidence of forgetting of the complete association between the Spanish word and the corresponding English word was found after a one-month delay. Subsequent experiments showed that with extended practice (over 50 repetitions) direct access of the English word was attained, and, thus, the speed of access was no longer influenced by the keyword used at the original learning. The retention test of these subjects' performance after a one-month delay showed, most interestingly, that recall of the English word ($p = 0.96$) was reliably higher than recall of the keyword ($p = 0.93$) when the Spanish word was presented as a cue. This finding suggests that at least some of the associations are no longer mediated, but also that the vast majority of keywords are still accessible. An analysis of the retrieval latencies showed no reliable difference at delay between recall of the English word and recall of the keyword and, compared to the latencies at the end of training, they were about 200 ms slower. A more detailed analysis looking separately at the performance on each of the two test blocks is currently under way.

With extended training, subjects were able to attain retrieval latencies for vocabulary items of around 800 ms. To estimate the lower bound of retrieval access, we examined the pilot data from bilingual speakers who were also given the task of retrieving the English word with the Spanish word as a cue. A preliminary analysis found that the bilingual subjects' initial retrieval took around 1,300 ms, but with a few repetitions, retrieval latencies of about 600 ms were attained. We have planned additional experiments to evaluate whether the observed speed-up reflects general processes and, thus, would transfer to vocabulary items not previously seen. Furthermore, we plan to examine the durability of this speed-up by administering retention tests at different delays.

In addition, we have developed a questionnaire which probes into our Spanish-speaking subjects' experience with both English and Spanish. The questionnaire asks subjects about the history of their language usage in both formal educational settings and less formal settings. For example, it asks if Spanish is a first or second language, if the subject has ever lived in a Spanish-speaking neighborhood or country, and how long the subject has studied Spanish. It also asks subjects to rate their knowledge/performance in speaking, understanding, and reading both Spanish and English. By developing such an extensive questionnaire we hope to identify functionally separable categories of language experience which map onto different levels of performance on the vocabulary task.

This project is being conducted by Robert Crutcher and Oliver Hammerle, with Anders Ericsson as the primary faculty advisor and Alice Healy as a secondary advisor.